

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-112835

(43)Date of publication of application : 28.04.1998

(51)Int. Cl.

H04N 5/76
G06F 17/30

(21)Application number : 08-264287

(71)Applicant : MATSUSHITA ELECTRIC IND CO
LTD

(22)Date of filing : 04.10.1996

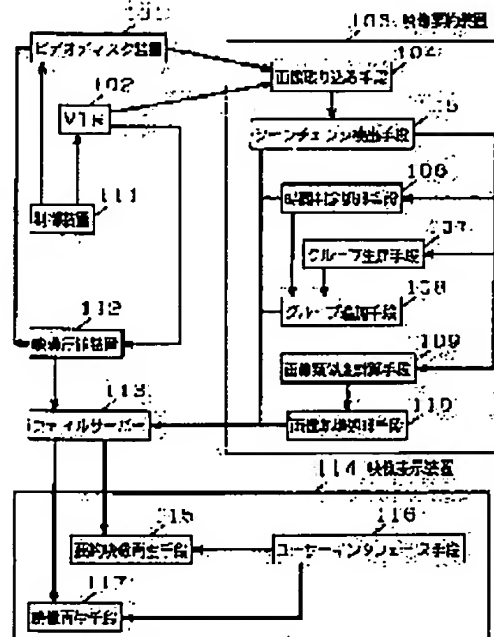
(72)Inventor : YAMADA SHIN
KIKUCHI YASUHIRO

(54) VIDEO IMAGE SUMMARIZING METHOD AND VIDEO IMAGE DISPLAY METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To allow the method to have provision for versatility of preference of the users and video image contents by extracting scenes consecutive for a prescribed time and grouping scenes in time series.

SOLUTION: When a scene change detection means 105 discriminates occurrence of a scene change in a frame image received by an image fetch means 104, the means decides a representative image in all scene modes and a head frame. Then a time length of a scene just before a scene change is calculated and a time discrimination processing means 106 stores a preceding scene representative image as a representative image in the time discrimination mode to a file server 113. Then the user selects in which mode a summarized video image is to be displayed via a user interface means 116, a head frame number of a scene to be displayed and information of the representative image are sent from the server 113 to a summarized video image reproduction means 115, which displays the summarized video image.



LEGAL STATUS

[Date of request for examination] 09.03.1999

[Date of sending the examiner's decision
of rejection][Kind of final disposal of application
other than the examiner's decision of
rejection or application converted
registration]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平10-112835

(43)公開日 平成10年(1998) 4月28日

(51)Int.Cl.⁴

H 0 4 N 5/76

G 0 6 F 17/30

識別記号

F I

H 0 4 N 5/76

G 0 6 F 15/40

15/401

B

5 3 0 Q

審査請求 未請求 請求項の数13 O L (全 17 頁)

(21)出願番号 特願平8-264287

(22)出願日 平成8年(1996)10月4日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)発明者 山田 伸

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

(72)発明者 菊池 廣弘

神奈川県川崎市多摩区東三田3丁目10番1

号 松下技研株式会社内

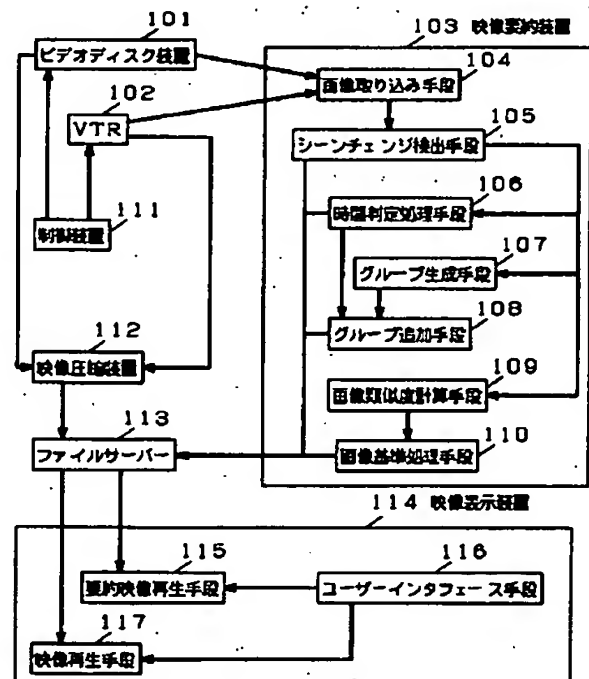
(74)代理人 弁理士 滝本 智之 (外1名)

(54)【発明の名称】 映像要約方法および映像表示方法

(57)【要約】

【課題】 映像内容の多様性および使用者の好みの多様性に対応する映像要約装置および要約情報を効率的に表示するための映像表示装置を提供することを目的とする。

【解決手段】 本発明の映像要約システムは取り込まれた映像を所定の基準に基づいて分割して形成した複数のシーンより映像の要約情報を抽出するための複数の映像要約手段を含む映像要約装置103 備えている。また、これらの要約情報を選択することのできる映像表示装置114 を備えている。この構成により、一定の映像要約基準のみで画一的に代表画像を選択する従来の映像要約装置や代表画像を映像の長さ等に関係なく選択する従来の映像表示装置と比べて、映像内容の多様性および使用者の好みの多様性に対応することができる。



【特許請求の範囲】

【請求項1】 取り込まれた映像をそのシーンチェンジを検出することにより分割して形成した複数のシーンについて、時系列的にその前後のシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）の類似度を計算することにより代表画像を含むシーンを時系列グループにまとめる時系列グループ生成過程と、同一の時系列グループ内にあっても一定時間以上継続するシーンについてはその前後のシーンとは独立する時系列グループとする修正をかける時系列グループ追加過程と、上記2つの過程で得られた各時系列グループの映像要約情報を出力する映像要約情報出力過程とからなる映像要約方法。

【請求項2】 時系列グループ生成過程におけるシーンの代表画像間の類似度を、比較すべき代表画像間で共通する色を持つ画素の割合を計算することを特徴とする請求項1に記載の映像要約方法。

【請求項3】 取り込まれた映像を構成する複数のシーンの内で所定の基準に従って選択された複数のシーン（以下、基準シーンという。）の所定の画像（以下、シーンの所定の画像を代表画像という。）と映像を構成する全てのシーン（以下、本編シーンという。）の代表画像との間で所定の基準に従って類似度を計算して、基準シーンの代表画像との類似度がしきい値以上になる代表画像を含むシーンを選択する類似度計算過程と、上記過程で得られたシーンの映像要約情報を出力する映像要約情報出力過程とからなる映像要約方法。

【請求項4】 類似度計算過程において、基準シーンの代表画像と本編シーンの代表画像との間の類似度を計算するための基準が、代表画像を複数の画像領域に分割し、上記両代表画像の各画像領域内の画素のRGB成分を比較するものであることを特徴とする請求項3に記載の映像要約方法。

【請求項5】 取り込まれた映像をそのシーンチェンジを検出することにより分割して形成した複数のシーンを複数の時系列グループにまとめるための複数の映像要約過程および/または上記複数のシーンより所定のシーンを選択するための複数の映像要約過程と、各映像要約過程により選択された各シーンの映像要約情報を出力する映像要約出力過程とからなる映像要約方法。

【請求項6】 複数のシーンを複数の時系列グループにまとめるための複数の映像要約過程および/または上記複数のシーンより所定のシーンを選択するための複数の映像要約過程が、全てのシーンを選択する第1の映像要約過程と、上記取り込まれた映像の複数のシーンのうち一定時間以上継続するシーンのみを選択する第2の映像要約過程と、時系列的にその前後のシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）の類似度を所定の基準に従って計算して類似度がしきい値以上になる代表画像を含むシーンを時系列グループにまと

める第3の映像要約過程と、請求項1または請求項2に記載した第4の映像要約過程と、請求項3または請求項4に記載した第5の映像要約過程のうち少なくとも2以上の映像要約過程であることを特徴とする請求項5に記載の映像要約方法。

【請求項7】 映像要約情報がシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）または所定のフレーム番号（以下、代表フレーム番号という。）であって、代表画像間のまたは代表フレーム番号間のフレーム数が多いものから順に代表画像または代表フレーム番号を選択して、所定の数のシーンまたは所定の数の時系列シーンを選択することを特徴とする、請求項5または請求項6に記載の映像要約方法

【請求項8】 一つの映像要約過程によっては所定の数の代表画像またはフレーム番号が選択できない場合に、残りの代表画像またはフレーム番号を他の映像要約手段によって選択された代表画像またはフレーム番号を選択することを特徴とする、請求項7に記載の映像要約方法。

【請求項9】 使用者が、映像要約過程の優先順位および選択すべきシーンまたは時系列シーンの数を入力し、その情報に従って映像要約を行うことを特徴とする請求項7または請求項8に記載の映像要約方法。

【請求項10】 映像および請求項1から請求項9までのいずれかの方法で選択した上記映像の代表画像を表示する映像表示方法において、代表画像の位置付近の映像の一部をつないだ映像（以下、要約映像という。）上で指定されたフレームを先頭にして、映像を再生することを特徴とする映像表示方法。

【請求項11】 映像の中から抜き出した画像の位置付近の映像の一部をつないだ映像（以下、要約映像という。）の再生の中断時に、中断した位置のフレーム画像の内容を代表する代表画像を含む複数の代表画像を一覧表示することを特徴とする映像表示方法。

【請求項12】 要約映像の先頭の一部が直前の要約映像の末尾の一部に含まれる場合には、当該要約映像の中で直前の要約映像の末尾の一部に含まれた最後のフレームの次のフレームを当該要約映像の先頭とすることを特徴とする請求項10または請求項11に記載の映像表示方法。

【請求項13】 取り込まれた映像を所定の基準に基づいて分割して形成した複数のシーンを複数の時系列グループにまとめることにより映像の要約情報を抽出するための複数の映像要約手段および/または上記複数のシーンより所定のシーンを選択することにより映像の要約情報を抽出するための複数の映像要約手段と、上記映像要約手段で抽出された要約情報を選択する要約情報選択手段および要約情報表示手段を備えた映像要約システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、映像の検索、編集、加工、早見などを支援する方法に係り、特にビデオテープやビデオディスクに格納された映像を要約して、再生または表示をする映像要約方法および映像表示方法に関するものである。

【0002】

【従来の技術】映像の編集、早見においては、映像の中から見たい部分を効率よく探すための映像要約手段が不可欠である。画像の中から特定の代表画像を選択することを特徴とする映像要約手段として、例えば、特開昭64-68084号に記載された方法(以下、シーソー表示という。)、ショット毎ラッシュ再生法、特開平7-236153号に記載された方法(以下、色差相関値法という。)、特開平6-149902号に記載された方法(以下、時間長指定要約法という。)が知られている。

【0003】シーソー表示は、シーンチェンジをあらかじめ調べておき、図14に示すように、シーンチェンジ直後の画像を一覧表示する方法である。図14において、(1)～(9)がシーンチェンジ直後の画像を表している。この方法では、一つのシーンをシーンの先頭画像で代表させる。

【0004】「シーン」とは、映像編集などの分野でしばしば使われる映像の単位であり、多くの場合、「一つのビデオカメラで時間的に連続して撮影された部分」と定義される。本願では「シーン」をこれより広い意味に用いるものとし、「シーン」とは、「映像を何らかの基準に従って分割して形成した映像の単位」をいうものとする。シーンは、ショット、カットと呼ばれることもある。「シーンチェンジ」とは、多くの場合、編集でつないだ部分、ビデオカメラの撮影を中断した部分のように、シーンが変化する部分をいう。本願ではこれより広い意味に用いるものとし、「シーンチェンジ」とは、「映像を何らかの基準に従って分割する際に分割された部分」をいうものとする。シーンチェンジは、カットと呼ばれることもある。また、「フレーム」とは、映像を構成する各画像をいうものとする。

【0005】映像を自動的にシーンに分割する方法としては、共通色比率法(山田、藤岡、金森、松島、「部分領域ごとの共通色に注目したシーンチェンジ検出方法の検討」、テレビジョン学会技術報告、Vol.17、No.55)、映像変化モデル法(山田、藤岡、金森、松島、坂内、「編集効果を含む映像のシーンチェンジ検出方法」、テレビジョン学会、マルチメディアと映像処理シンポジウム'94)などが提案されている。

【0006】ショット毎ラッシュ再生法は、各シーンの先頭部分を標準速度で次々に再生する映像要約方法である。この方法は、映像の要約を見るための方法であり、途中で再生を中断せずに最初から最後まで再生して利用する。

【0007】時間長指定要約法は、全部または一部のシーンを一定時間づつ再生する方法であり、要約動画の時間長が指定された値になるように、各シーンの時間長を用いて、各シーンの再生時間または再生シーン数を決める。

【0008】色差相関値法は、カメラアングルのみが変わって、同様な内容をもつシーンをグループ化することで、シーソー表示に比べて能率的な映像検索を可能にする映像要約方法である。この方法では、シーンの先頭画像(以下、カット画面という。)を求めながら、カット画面間の色差ヒストグラム相関値を求め、色差ヒストグラム相関値をしきい値以上にするカット画面を同一グループのカット画面とみなす。このとき、類似背景を有するカット画面同士が、同一グループになる。また、カット画面を一覧表示する場合には、同一グループ内で最初に検出されたカット画面を親画面にして、同一グループ内の残りのカット画面を子画面にして表示する。例えば、2番目～3番目のシーンが同一グループで、6番目～8番目のシーンが同一グループであるとき、図15に示すように表示する。図15において、(1)～(9)は、それぞれ1番目のカット画面から9番目のカット画面までを表す。

【0009】以下、色差相関値法を用いた従来の映像要約システムについて説明する。図16は従来例である映像要約システムを示すブロック図である。図16において、1501、1502は映像の入力装置であって、1501はビデオディスク装置、1502はVTRである。1503はビデオディスク装置1501やVTR1502からの映像信号を処理しながら、シーンチェンジ検出と色差相関値法を用いて映像を要約する映像要約装置であり、フレーム画像を取り込む画像取り込み手段1504と、シーンチェンジを検出するシーンチェンジ検出手段1505と、カット画面(シーンの先頭画像)のグループ化を実行するグループ生成手段1506から構成される。1507は、ビデオディスク装置1501とVTR1502を制御する制御装置である。1508は、映像を圧縮する映像圧縮装置である。1509は映像要約装置1503で作成されたカット画面とシーンチェンジのデータとグループのデータと、映像圧縮装置1508で圧縮された映像データを保存するファイルサーバーである。1510はファイルサーバーに格納されたデータと画像と映像を表示する映像表示装置である。

【0010】以上のように構成された映像要約システムについて、図17に示すフローチャートを用いてその全体の動作を説明する。

【0011】手順1601では、図16における制御装置1507がビデオディスク装置1501とVTR1502を制御して、映像の再生を開始し、同時に映像圧縮装置1508での映像の圧縮を開始する。圧縮された映像はファイルサーバー1509に保存する。

5

【0012】手順1602では、制御装置1507が映像が終了したかどうか判定する。映像が終了した場合には手順1607に進み、そうでなければ、手順1603に進む。

【0013】手順1603では、画像取り込み手段1504が再生中のフレーム画像を取り込む。

【0014】手順1604では、シーンチェンジ検出手段1505が画像取り込み手段1504で取り込まれたフレーム画像を処理して、前述した共通色比率法等を用いてシーンチェンジが発生したかどうか検出する。

【0015】手順1604で「シーンチェンジが発生した」と判定された場合には手順1605に進み、そうでなければ手順1602に戻る。

【0016】手順1605では、シーンチェンジ検出手段1505で検出したシーンチェンジが発生した時点のフレーム番号とカット画面をファイルサーバ1509に保存する。

【0017】手順1606では、グループ生成手段1506が、時系列のカット画面間の色差ヒストグラム相関値を求め、色差ヒストグラム相関値をしきい値以上にするカット画面を同一グループのカット画面とみなす。グループ化の結果をファイルサーバ1509に保存してから、手順1602に戻る。

【0018】手順1607では、制御装置1507が映像の再生と映像の圧縮を中止する。手順1608では、使用者が映像の表示方法を選択する。代表的な部分を動画で見たい場合には、手順1610に進む。静止画を用いて見たい部分を探したい場合には、手順1609に進む。

【0019】手順1609では、使用者が見たい部分を効率よく探せるように、システムが映像表示装置1510上に映像の要約を表示する。例えば、図14に示したように、カット画面を一覧表示する。また、同一グループと判定されたカット画面を、同一グループ内で最初に検出されたカット画面と連結して子画面の形式で表示してもよい。2番目のシーンと3番目のシーンが同一グループで、6番目～8番目のシーンが同一グループであるとき、図15のように表示されることになる。

【0020】手順1610では、システムが映像表示装置1510上に映像の各グループの先頭部分を所定の時間ずつ、例えば5秒ずつ表示する。

【0021】

【発明が解決しようとする課題】従来、画像の中から特定の代表画像を選択することを特徴とする映像要約装置は、一つの番組の中でどの映像が重要であるかがその映像の内容等により異なるにもかかわらず、一つの基準のみにより代表画像を選択していた。例えば「従来の技術」の欄で説明した色差相関値法を用いた映像要約装置では、「カメラアングルのみが異なる同様な内容をもつシーンは同一のグループである」という映像要約基準の

6

みによって番組の代表画像を選択していた。しかし同方法では、例えばニュース番組においてしばしば起こるように、現場のアナウンサーが事件を説明するシーン等のように同一背景であっても人間が現れているシーンやその直後のシーンのように重要な情報を有している可能性が高いシーンを代表画像として選択することができない。これらの画像は背景が類似している限り全てまとめられてしまうからである。また同方法では、衛星放送の番組「ハイテクシャワーインターナショナル」等のように本編の各記事から2～3シーンの動画像を抜き出して作成されたダイジェストが最初に流れてから本編が流れる番組で、これらのダイジェスト画像を代表画像となるように本編をグループ化することができない。ダイジェスト部分と本編部分とを区別せず、隣り合うシーン間のみで類似度を判断するからである。

【0022】映像要約装置のみならず、その映像要約装置により得られた要約情報を含む映像情報を表示する映像表示装置についても課題があった。すなわち、全てのシーンの一部又は全部を何らかの方法で再生するショット毎ラッシュ再生法や時間長指定要約法に基づく映像表示装置においては、映像のシーンの数が多いときには代表画像も多くなり、使用者側の負担となっていた。また、画像の中から所定の基準に従って代表画像を選択する映像要約装置により得られた要約情報を含む映像情報を表示する映像表示装置においても、映像のシーンの数が多いときには代表画像が多くなることが多く、その結果、一覧表示される画像が多くなり、見たい部分を探しにくいということがしばしばあった。例えばニュース番組を要約した場合には、多くの場合、各記事から数十枚のシーンの代表画像が表示されるために、一つの記事を見終わるまでに、数十枚のシーンの代表画像を見る必要があるという課題があった。

【0023】また、図16、図17を用いて説明したシステムにおける映像表示装置では静止画を用いて見たい部分を探すことはできるが、動画と音声を用いて見たい部分を探すことができないという課題を有していた。さらに、ショット毎ラッシュ再生法と時間長指定要約法では静止画を用いて見たい部分を探すことができないだけでなく、一定時間に達しないシーンが再生される場合、その一定時間の中に次のシーンの先頭の映像が含まれるのでその一定時間の終わりまで次のシーンが再生され、その後再び当該次のシーンが先頭から一定時間再生されるので、この当該次のシーンの先頭の映像が連続してかつ重複して再生されるという課題を有していた。

【0024】以上で述べたように、映像の内容や使用者の好み等に関わりなく一定の映像要約基準のみで画一的に代表画像を選択する映像要約装置や代表画像を映像の長さ等に関係なく表示する映像表示装置では、映像の内容等に依存する多様性および要約に対する使用者の多様なニーズに対応できないという課題を有していた。

10

20

30

40

50

【0025】本発明は上記従来技術の課題を解決するもので、映像内容の多様性および使用者の好みの多様性に対応する映像要約装置および要約情報を効率的に表示するための映像表示装置を提供することを目的とする。

【0026】

【課題を解決するための手段】上記課題を解決するために、本発明に係る映像要約方法を実現する映像要約装置は、取り込まれた映像を所定の基準に基づいて分割して形成した複数のシーンを複数の時系列グループにまとめるためのまたは上記複数のシーンより所定のシーンを選択するための映像要約手段を複数用意している。この複数の映像要約手段とは、一定時間以上継続するシーンだけを検出してその代表画像を抜き出す手段、色差相関値法などを用いてグループ化した結果に対して一定時間以上継続するシーンがその前後のシーンと同一グループにならないように修正してから各グループの代表画像を抜き出す手段、さらに、映像のダイジェストを含む数十シーンのそれぞれの代表画像と他のシーンのそれぞれの代表画像との間で類似度を計算し類似度がしきい値以上になるシーンを検出してその代表画像だけを抜き出す手段等であり、各手段によって抜き出された映像要約情報は内部または外部に備えられた記録手段に映像情報自体と共に記録される。本発明に係る映像表示方法を実現する映像表示装置では、上記記録手段に記録された映像要約情報である映像の要約画像である代表画像を一覧表示し、一覧表示された画像の中から見たい部分の画像を指定すると、その部分以降の映像が再生される。さらに、あらかじめ映像要約手段の優先順位と、各映像要約手段の中での画像選択の優先順位を決めておき、これらの優先順位を用いて代表画像を決定し、使用者に指定された枚数以下の画像を抜き出すことができる。また、一つの映像要約手段により抜き出された代表画像の枚数がしきい値以上になると、当該映像要約手段の代表画像を他の映像要約手段による代表画像よりも先に選んで映像から抜き出す画像を決定し、そうでない場合には、当該映像要約手段以外の他の映像要約手段の代表画像の中から選んで映像から抜き出す画像を決定して、使用者に指定された枚数以下の画像を抜き出すことができる。

【0027】さらに、代表画像を一覧表示する代わりに、各代表画像の位置付近の映像の一部を次々に再生する（以下、再生された映像を要約映像という）ことができる。そして、この要約映像上の任意のフレームを先頭にして、要約前の映像を再生することができる。また、要約映像の再生の中断時に、中断した位置のフレーム画像の内容を代表する代表画像を含む複数の代表画像を一覧表示する。

【0028】ただし、要約映像において、代表画像の位置付近の映像の一部が直前の代表画像の位置付近の映像の一部に含まれる場合には、直前の代表画像の位置付近の映像の一部の末尾を、代表画像の位置付近の映像の一

部の先頭とする。

【0029】

【発明の実施の形態】請求項1に記載の発明は、取り込まれた映像をそのシーンチェンジを検出することにより分割して形成した複数のシーンについて、時系列的にその前後のシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）の類似度を計算することにより代表画像を含むシーンを時系列グループにまとめる時系列グループ生成過程と、同一の時系列グループ内にあっても一定時間以上継続するシーンについてはその前後のシーンとは独立する時系列グループとする修正をかける時系列グループ追加過程と、上記2つの過程で得られた各時系列グループの映像要約情報を出力する映像要約情報出力過程とからなる映像要約方法であり、一定時間以上継続するシーンは全て抜き出すとともに、所定の基準で時系列のシーンをグループ化するという作用を有する。

【0030】請求項2に記載の発明は、請求項1に記載した映像要約方法の時系列グループ生成過程におけるシーンの代表画像間の類似度を計算する方法が、比較すべき代表画像間で共通する色を持つ画素の割合を計算するものであることを特徴とする映像要約方法であり、一定時間以上継続するシーンは全て抜き出すとともに、色に関する類似度を基準として時系列のシーンをグループ化するという作用を有する。

【0031】請求項3に記載の発明は、取り込まれた映像を構成する複数のシーンの内で所定の基準に従って選択された複数のシーン（以下、基準シーンという。）の所定の画像（以下、シーンの所定の画像を代表画像という。）と映像を構成する全てのシーンの代表画像との間で所定の基準に従って類似度を計算して基準シーンの代表画像との類似度がしきい値以上になる代表画像を含むシーンを選択する類似度計算過程と、上記過程で得られたシーンの映像要約情報を出力する映像要約情報出力過程とからなる映像要約方法であり、映像のダイジェストシーンである蓋然性の高い時系列シーンと映像を構成する全てのシーンとの間の類似度を求めることによりダイジェストシーンに対応する本編シーンを抜き出すという作用を有する。

【0032】請求項4に記載の発明は、請求項3に記載の映像要約方法であって、類似度計算過程において基準シーンの代表画像と本編シーンの代表画像との間の類似度を計算するための基準が、代表画像を複数の画像領域に分割し、両代表画像の各画像領域内の画素の平均色のRGB成分を比較するものであることを特徴とする。映像のダイジェストシーンである蓋然性の高い時系列シーンと映像を構成する全てのシーンとの間の代表画像の各画像領域内の画素の平均色のRGB成分を比較することにより類似度を求めることで、ダイジェストシーンに対応する本編シーンを抜き出すという作用を有する。

【0033】請求項5に記載の発明は、取り込まれた映像をそのシーンチェンジを検出することにより分割して形成した複数のシーンを複数の時系列グループにまとめるための複数の映像要約過程および／または上記複数のシーンより所定のシーンを選択するための複数の映像要約過程と、各映像要約過程により選択された各シーンの映像要約情報を出力する映像要約出力過程とからなる映像要約方法であって、使用者が映像の特徴等に合わせて代表画像を決定するための複数の映像要約情報を用意するという作用を有する。

【0034】請求項6に記載の発明は、請求項5に記載した映像要約方法であって、複数のシーンを複数の時系列グループにまとめるための複数の映像要約過程および／または上記複数のシーンより所定のシーンを選択するための複数の映像要約過程が、全てのシーンを選択する映像要約過程と、上記取り込まれた映像の複数のシーンのうち一定時間以上継続するシーンのみを選択する映像要約過程と、時系列的にその前後のシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）の類似度を所定の基準に従って計算して類似度がしきい値以上になる代表画像を含むシーンを時系列グループにまとめる映像要約過程と、請求項1または請求項2に記載した映像要約過程と、請求項3または請求項4に記載した映像要約過程のうち少なくとも2以上の映像要約過程であることを特徴とするものであり、使用者が映像の特徴等に合わせて代表画像を決定するための複数の映像要約情報を用意するという作用を有する。

【0035】請求項7に記載の発明は、請求項5または請求項6に記載した映像要約方法であって、映像要約情報がシーンの所定の画像（以下、シーンの所定の画像を代表画像という。）または所定のフレーム番号（以下、代表フレーム番号という。）であって、代表画像間のまたは代表フレーム番号間のフレーム数が多いものから順に代表画像または代表フレーム番号を選択してゆき、所定の数のシーンまたは時系列シーンを選択するという作用を有する。

【0036】請求項8に記載の発明は、請求項7に記載の映像要約方法であって、一つの映像要約過程によっては所定の数のシーンまたは時系列シーンが選択できない場合に、残りのシーンまたは時系列シーンを他の映像要約手段によって選択されたシーンまたは時系列シーンを選択するという作用を有する。

【0037】請求項9に記載の発明は、請求項7または請求項8に記載の映像要約方法であって、使用者が、映像要約過程の優先順位および選択すべきシーンまたは時系列シーンの数を入力し、その情報に従って映像要約を行うという作用を有する。

【0038】請求項10に記載の発明は、映像および請求項1から請求項9までのいずれかの方法で選択した上記映像の代表画像を表示する映像表示方法において、代

表画像の位置付近の映像の一部をつないだ映像（以下、要約映像という。）上で指定されたフレームを先頭に置いて、映像を再生することを特徴とする映像表示方法であって、音声や被写体の動きを考慮しながら見たい部分を指定できるという作用を有する。

【0039】請求項11に記載の発明は、映像の中から抜き出した画像の位置付近の映像の一部をつないだ映像（以下、要約映像という。）の再生の中断時に、中断した位置のフレーム画像の内容を代表する代表画像を含む複数の代表画像を一覧表示することを特徴とする映像表示方法であって、使用者が見たい部分の情報を集中的に表示するという作用を有する。

【0040】請求項12に記載の発明は、請求項10または請求項11に記載の映像表示方法であって、要約映像の先頭の一部が直前の要約映像の末尾の一部に含まれる場合には、当該要約映像の中で直前の要約映像の末尾の一部に含まれた最後のフレームの次のフレームを当該要約映像の先頭とするという作用を有する。この方法は一度再生された要約映像を間髪入れずに重複しての再生を回避することができる。

【0041】請求項13に記載された発明は、取り込まれた映像を所定の基準に基づいて分割して形成した複数のシーンを複数の時系列グループにまとめることにより映像の要約情報を抽出するための複数の映像要約手段および／または上記複数のシーンより所定のシーンを選択することにより映像の要約情報を抽出するための複数の映像要約手段と、上記の一またはそれ以上の映像要約手段で抽出された要約情報を選択する要約情報選択手段および要約情報表示手段を備えた映像要約システムであって、どの基準で要約された情報を表示するかを自由に選択できるという作用を有する。

【0042】以下、本発明に基づく実施の形態を図面を参照しながら説明する。

（第1の実施の形態）図1は第1の実施の形態であって、複数の映像要約基準に基づき映像より要約情報を抽出する映像要約装置と、この装置により選択された映像の要約情報およびその映像自体を表示する映像表示装置を組み合わせ構成した映像要約システムを示すブロック図である。

【0043】図1において、101、102は映像出力装置であって、101はビデオディスク装置、102はVTRである。103はビデオディスク装置101やVTR102からの映像信号を処理しながら、映像を要約する映像要約装置であって、フレーム画像を取り込む画像取り込み手段104と、シーンチェンジを検出するシーンチェンジ検出手段105と、一定時間以上継続するシーンを検出する時間判定処理手段106と、シーンの代表画像のグループ化を実行するグループ生成手段107と、時間判定処理手段106とグループ生成手段107の出力を受けてグループ化の結果を修正するグループ

10

20

30

40

50

追加手段108と、シーンの代表画像の間の類似度を計算する画像類似度計算手段109と、類似度計算手段で求めた類似度がしきい値以上になるシーンを検出する画像基準処理手段110から構成される。111は、ビデオディスク装置101とVTR102を制御する制御装置である。112は、映像を圧縮する映像圧縮装置である。113は映像圧縮装置112で圧縮された映像データと、映像要約装置103で検出された各種シーンとグループのデータと、それぞれの代表画像を保存するファイルサーバーである。114はファイルサーバ113に格納されたデータを用いて映像を表示する映像表示装置であり、映像要約装置で抜き出された画像の位置付近の映像を一定時間づつ徐々に再生する要約映像再生手段115と、映像の再生を制御するユーザーインターフェース手段116と、指定された位置以降の映像を再生する映像再生手段117から構成される。

【0044】映像要約システムの映像要約装置103は、例えば、コンピュータハードウェアおよびソフトウェアの組み合わせにより実現することができる。また、映像表示装置114のうち、要約映像再生手段115および映像再生手段117はコンピュータ上で実現することができ、ユーザーインターフェース手段116はCRT等のモニターおよびコンピュータのハードウェアおよびソフトウェアの組み合わせにより実現することができる。

【0045】本実施の形態に係る映像要約装置103は4つの映像要約基準を備えている。第1の映像要約基準は、映像の含む全てのシーンの先頭画面を代表画像とする映像要約基準である。以下では、この映像要約基準を全シーン表示モードという。第2の映像要約基準は、一定値以上の時間長を有するシーンの先頭画面を代表画像とする映像要約基準である。以下では、この映像要約基準を時間判定モードという。第3の映像要約基準は、時間判定モードで選択されたシーンを除くシーンを色差相関値法によりグループ化し、時間判定モードで選択されたシーンおよび上記グループ化されたシーンの先頭画面を代表画像とする映像要約基準である。以下では、この映像要約基準をハイブリッドモードという。第4の映像要約基準は、特定のニュース番組のように映像のダイジェストが番組の冒頭等にあることがあらかじめわかっている映像の当該ダイジェストを含む部分の代表画像とその他のシーンの先頭画面との間の類似度を、複数の画像領域内の画素のRGB成分の比較によって計算し、所定のしきい値以上の値を持つシーンの先頭画面を代表画像とする映像要約基準である。以下では、この映像要約基準を画像基準モードという。

【0046】以上のように構成される映像要約システムについて、図2に示すフローチャートを用いてその動作を説明する。

【0047】手順201では、図1における制御装置1

11がビデオディスク装置101とVTR102を制御して、映像の再生を開始し、同時に映像圧縮装置112での映像の圧縮を開始する。

【0048】手順202では、制御装置111が映像が終了したかどうか判定する。映像が終了した場合には手順212に進み、そうでなければ、手順203に進む。

【0049】手順203から手順211までは映像要約装置103において行われる処理である。

【0050】手順203では、画像取り込み手段104が再生中のフレーム画像を取り込む。

【0051】手順204では、シーンチェンジ検出手段105が、画像取り込み手段104で取り込まれたフレーム画像を処理して、「従来の技術」の欄で述べた共通色比率法を用いて、シーンチェンジが発生したかどうかを判定する。ただし、共通色比率法では、次シーンのフレーム画像を2枚処理してからシーンチェンジが検出されるので、「シーンチェンジが発生した」という判定は「前回取り込んだフレーム画像がカット画面である」とことを表す。

【0052】図3は、今述べたシーンチェンジ検出手段105がシーンチェンジを検出する方法と前シーンを示す説明図である。図3において、aは手順203で取り込まれたフレーム画像であり、bは1回前の手順203で取り込んだフレーム画像であり、cは2回前の手順203で取り込んだフレーム画像であり、dは3回前の手順203で取り込んだフレーム画像である。この場合、シーンチェンジはbとcとの間で発生しているが、それを検出するのはaを取り込んだ後であることを示している。また、検出したシーンチェンジの直前のシーンを前シーンと呼ぶ。

【0053】なお、映像変化モデル法などを用いてシーンチェンジを検出してもよい。また、あらかじめオペレータがシーンチェンジを判断して、そのシーンチェンジが発生した時点のフレーム番号を入力しておいてもよい。

【0054】手順204でシーンチェンジ検出手段105が「シーンチェンジが発生した」と判定した場合には手順205に進み、そうでなければ手順202に戻る。

【0055】手順205では、全シーンモードのシーンの代表画面および先頭フレームを決定する。具体的には、手順204で取り込まれた画像およびシーンチェンジが発生した時点のフレーム番号を全シーンモードのシーンの代表画面および先頭フレームとしてファイルサーバ113に保存する。なお、画像をファイルサーバ113に保存する場合には、縮小して保存してもよい。

【0056】手順206では、時間判定モードのシーンの代表画面および先頭フレームを決定する。具体的には、手順204で検出したシーンチェンジの直前のシーン（以下、前シーンという）の時間長を計算してから、時間判定処理手段106で、時間長が8秒以上になる前

10

20

30

40

50

13

シーンを検出し、検出した前シーンの代表画像を時間判定モードの代表画像とみなしてファイルサーバ113に保存する。また、時間判定処理手段106で検出された前シーンの先頭フレーム番号を時間判定モードで表示するシーンの先頭フレーム番号としてファイルサーバに保存する。

【0057】なお、本実施の形態では、8秒以上の前シーンを検出したが、必ずしも8秒である必要はない。

【0058】手順207から209まででは、ハイブリッドモードのシーンの代表画面および先頭フレームを決定する。

【0059】手順207では、グループ生成手段107が、手順204で検出したシーンチェンジ直後のシーン（以下、次シーンという。）の代表画像と前シーンの代表画像をファイルサーバ113から取り出し、その間の色差ヒストグラム相関値を求め、色差ヒストグラム相関値がしきい値以上になる場合に、次シーンの代表画像をハイブリッドモードの代表画像とみなしてファイルサーバ113に保存する。さらに、色差ヒストグラム相関値がしきい値以上になる場合には、次シーンの先頭フレーム番号をハイブリッドモードで表示するシーンの先頭フレーム番号としてファイルサーバ113に保存する。

【0060】手順208は分岐処理である。前シーンの時間長が8秒以上になる場合には、手順209に進み、そうでなければ手順210に進む。

$$S(M, N) = \sum_{i=1}^{10} (|R_i(M) - R_i(N)| + |G_i(M) - G_i(N)| + |B_i(M) - B_i(N)|) \cdot \alpha \quad (1)$$

【0065】（数1）式において、 $|x|$ は x の絶対値を表す。また、従来例で述べた色差ヒストグラム相関値を類似度として用いてもよい。

【0066】なお、以上の説明では、次シーンの代表画像と間の類似度を計算する画像として、映像の先頭の20シーンの代表画像を用いたが、これは、番組の構成によって変更するようにすることもできる。具体的には、映像表示装置114のユーザインターフェース手段116からの命令により画像類似度計算手段109が次シーンの代表画像 I_N を映像の先頭でない部分から選ぶようにしてもよい。例えば、映像の末尾にダイジェストがある場合には、映像の末尾の複数シーンをを用いる。

【0067】手順211では、画像基準処理手段110で、 M が1から20までのいずれかの値をとるときに、

【0068】

【数2】 $S(M, N) > \theta_{SIM} \dots \dots (2)$

【0069】が成り立つかどうか調べ、（数2）が成立するときに、次シーンの先頭フレーム番号を画像基準モードで表示するシーンの先頭フレーム番号としてファイルサーバ113に保存し、次シーンの代表画像を画像基準モードの代表画像とみなす。（数2）において θ

14

*【0061】手順209では、グループ追加手段108が時間長が8秒以上の前シーンの代表画像と次シーンの代表画像をハイブリッドモードの代表画像とみなしてファイルサーバ113に保存する。さらに、上記前シーンの先頭フレーム番号と次シーンの先頭フレーム番号をハイブリッドモードで表示するシーンの先頭フレーム番号としてファイルサーバ113に保存する。

【0062】手順210から211まででは、画像基準モードのシーンの代表画面および先頭フレームを決定する。

【0063】手順210では、手順204においてシーンチェンジ検出手段105により求められた次シーンの代表画像 I_N と映像の先頭の20シーンの代表画像 I_M （ M は1から20までの自然数）との間の類似度 $S(M, N)$ を画像類似度計算手段110が計算する。類似度の計算方法としては、各種の方法が考えられる。図4でその一例を紹介する。図4は画像基準モードにおいて、ダイジェストを含む部分の代表画像とその他のシーンとの間の色彩の類似度を計算する際に画面をブロックへの分割することを示す説明図である。すなわち、代表画像 I_M を図4に示すように4×4個のブロックに分割し、 i 番目のブロック内の画素の平均色のRGB成分 $R_i(M)$ 、 $G_i(M)$ 、 $B_i(M)$ を用いて次式で計算すればよい。

【0064】

【数1】

SIM はあらかじめ設定したしきい値である。

【0070】手順211の終了後、手順202に戻る。手順212では、映像の再生と映像の圧縮を中止する。

【0071】ここまでの手順でファイルサーバ113には、複数の映像要約基準、すなわち全シーンモード、時間判定モード、ハイブリッドモード、画像基準モードによりそれぞれ抜き出された要約情報、すなわち代表画像と先頭フレーム番号が保存された。以下の手順213から手順215までは表示装置114において行われる処理である。

【0072】手順213では、使用者がユーザインターフェース手段116を介してどのモードによる要約映像を表示するかを選択する。選択肢は、上述したように、時間判定モード、ハイブリッドモード、画像基準モード、全シーン表示モードである。ユーザインターフェース手段116は、ファイルサーバ113に対して、選択されたモードのシーンの代表画像および先頭フレームの情報を要約映像再生手段115に送る旨の信号を出力する。

【0073】手順214では、手順213で選択されたモードで表示されるべきシーンの先頭フレーム番号と代表画像の情報が、ファイルサーバ113より要約映像

再生手段115に送られる。これらの情報に基づいて、要約映像再生手段115は要約映像を表示する。ここでは使用者がハイブリッドモードを選択した場合について説明する。図5は本発明の第1の実施の形態における要約映像の作成方法を示す説明図である。3段の映像情報のうち、上段は映像自体、中段は上段の映像より抜き出された5秒間の要約映像、下段は中段の要約映像のみを連続的に再生することを示したものである。すなわち、図5の下段に示すように、ハイブリッドモードの各代表画像のシーンの先頭以降の5秒間の映像データをファイルサーバー113から取り出し、標準速度で次々に再生する。

【0074】なお、各代表画像のシーンの先頭以降の映像データの長さは5秒間でなくてもよい。また、各代表画像のシーンの先頭以降の5秒間ではなく、「各代表画像の2秒前から5秒間再生する」というように、各代表画像の位置前後の映像を次々に再生してもよい。また、標準速度で再生せずに、早送りで再生してもよい。また、映像内容に合わせて早送りの速度を変えてもよい。また、各代表画像のシーンの先頭以降の5秒間が複数のシーンを含んでもよい。

【0075】使用者は、ユーザーインタフェース手段116を用いて要約映像の表示をしながら、見たい部分を探す。要約映像の制御ボタンとしては、再生、逆再生、静止、早送り、巻き戻し、1コマ送り、1コマ戻しなどが考えられる。

【0076】手順215では、使用者が要約映像を見ながら見たい部分の先頭を指定する。図6は、要約映像の再生部分を示す説明図である。すなわち、使用者がi番目のシーンの代表画像の先頭を2秒見た時点でユーザーインタフェース手段116によりそのシーンを指定した場合、映像再生手段117による映像の再生開始時点は「i番目のシーンの代表画像の先頭から2秒経過した時点」となる。

【0077】図7に、見たい部分を指定するためのユーザーインターフェース116の画面の例を示す。図7において、aは要約映像を表示する部分、bは要約映像を制御するボタン、cは見たい部分を指定するボタンである。見たい部分を指定するボタンを押した瞬間に、aに表示されている画像が、見たい部分の先頭とみなされる。従って、見たい部分で要約映像を静止させておき、ボタンcを押すと、確実に見たい部分の先頭を指定することができる。

【0078】なお、以上の説明では、グループの先頭シーンの代表画像をハイブリッドモードの代表画像としたが、他の方法で代表画像を選んでもよい。例えば、各グループで最も時間長が長いシーンの代表画像をハイブリッドモードの代表画像とみなしてもよい。また、本実施の形態では選択するモードを4つ用意したが、4つでなくともよい。代表画像を求める順番も自由であり、必ず

しも、全シーン表示モード、時間判定モード、ハイブリッドモード、画像基準モードの順番で代表画像を決定しなくてもよい。

【0079】手順214において、要約映像を静止したときに、静止した位置の直前の代表画像の前後の代表画像を一覧表示してもよい。そして、手順215において、表示された代表画像の一つを見たい部分の先頭として指定してもよい。同時に表示する代表画像の枚数は、例えば24枚にすればよい。

【0080】また、手順214において、各代表画像のシーンの先頭以降の5秒間の映像データを取り出すが、直前の代表画像のシーンの先頭以降の5秒間の中に代表画像のシーンの先頭フレームが含まれることがある。図8は、かかる状況での要約映像の再生部分を示す説明図である。上記説明のように、取り出した映像データを次々に再生すると、図8のAの部分が2回再生されることになる。そこで、図示したように、「直前の代表画像のシーンの先頭以降の5秒間」の中に「代表画像のシーンの先頭フレーム」が含まれる場合には、代表画像のシーンの先頭ではなく、直前に再生した部分（例えば、図8のAの部分）の直後から再生してもよい。

【0081】（第2の実施の形態）図9は第2の実施の形態であって、複数の映像要約基準に基づき映像より要約情報を抽出するとともに使用者に指定された枚数以下の代表画像を抜き出す映像要約装置と、この装置により選択された映像の要約情報およびその映像自体を表示する映像表示装置を組み合わせる構成した映像要約システムを示すブロック図である。

【0082】図9において、801、802は映像出力装置であって、801はビデオディスク装置、802はVTRである。803はビデオディスク装置801やVTR802からの映像信号を処理しながら、使用者に指定された枚数以下の画像を抜き出して映像を要約する映像要約装置であって、フレーム画像を取り込む画像取り込み手段804と、シーンチェンジを検出するシーンチェンジ検出手段805と、一定時間以上継続するシーンを検出する時間判定処理手段806と、シーンの代表画像のグループ化を実行するグループ生成手段807と、時間判定処理手段806とグループ生成手段807の出力を受けてグループ化の結果を修正するグループ追加手段808と、シーンの代表画像の間の類似度を計算する画像類似度計算手段809と、類似度計算手段で求めた類似度がしきい値以上になるシーンを検出する画像基準処理手段810と、シーンチェンジ検出手段805と時間判定処理手段806とグループ追加手段808と画像基準処理手段810の出力を受けて代表画像を決定する画像判定手段811から構成される。812は、ビデオディスク装置801とVTR802を制御する制御装置である。813は、映像を圧縮する映像圧縮装置である。814は映像要約装置803で決定した代表画像と

そのフレーム番号を保存するファイルサーバーである。815はファイルサーバー814に格納されたデータと画像と映像を表示する映像表示装置である。

【0083】映像要約システムの映像要約装置803は、例えば、コンピュータ上で実現することができる。また、映像表示装置815は、例えば、コンピュータと、CRT等とモニターの組み合わせによる実現することができる。

【0084】本実施の形態に係る映像要約装置は4つの映像要約基準を備えている。それらは第1の実施の形態において説明したものと同じであるのでその説明は省略する。

【0085】以上のように構成された映像要約システムについて、図10に示すフローチャートを用いてその動作を説明する。

【0086】手順901では、手順201と同様に、図9における制御装置812がビデオディスク装置801とVTR802を制御して、映像の再生を開始し、同時に映像圧縮装置813での映像の圧縮を開始する。

【0087】手順902では、手順202と同様に、映像が終了したかどうか判定する。映像が終了した場合には手順912に進み、そうでなければ、手順903に進む。

【0088】手順903では、手順203と同様に、画像取り込み手段804が再生中のフレーム画像を取り込む。

【0089】手順904では、手順204と同様に、シーンチェンジ検出手段805が、画像取り込み手段804で取り込まれたフレーム画像を処理して、共通色比率法等を用いてシーンチェンジが発生したかどうかを判定する。

【0090】手順904で「シーンチェンジが発生した」と判定された場合には手順905に進み、そうでなければ手順902に戻る。

【0091】手順905では全シーンモードのシーンの代表画面および先頭フレームを決定する。すなわち、手順904で取り込まれた画像およびその画像のフレーム番号をファイルサーバー814に保存する。保存した画像は、シーンの代表画像として用いる。また、シーンの代表画像を全シーン表示モードの代表画像とみなす。

【0092】手順906では時間判定モードのシーンの代表画面および先頭フレームを決定する。すなわち、手順206と同様に、手順904で検出したシーンチェンジの直前のシーン（以下、前シーンという）の時間長を計算してから、時間判定処理手段806で、時間長が8秒以上になる前シーンを検出し、検出した前シーンの代表画像を時間判定モードの代表画像とみなす。

【0093】手順907から手順909まででは、ハイブリッドモードのシーンの代表画面および先頭フレームを決定する。

【0094】手順907では、手順207と同様に、グループ生成手段807が、手順904で検出したシーンチェンジ直後のシーン（以下、次シーンという。）と前シーンの代表画像間の色差ヒストグラム相関値を求め、色差ヒストグラム相関値がしきい値以上になる場合に、次シーンの代表画像をハイブリッドモードの代表画像とみなしてファイルサーバー814に保存する。なお、色差ヒストグラムを用いずに、共通画素法を用いて時系列のシーンのグループ化を実行し、グループの先頭シーンの代表画像をハイブリッドモードの代表画像とみなしてもよい。共通画素法については、本実施の形態に係る映像要約装置の動作を説明した後で説明する。

【0095】手順908では、手順208と同様に分岐処理を実行する。前シーンの時間長が8秒以上になる場合には、手順909に進み、そうでなければ手順910に進む。

【0096】手順909では、手順209と同様に、グループ追加手段808が、前シーンの代表画像と次シーンの代表画像をハイブリッドモードの代表画像とみなしてファイルサーバー814に保存する。

【0097】手順210から手順211まででは、画像基準モードのシーンの代表画面および先頭フレームを決定する。

【0098】手順910では、手順210と同様に、次シーンの代表画像 I_n と映像の先頭の20シーンの代表画像 I_m (M は1から20までの自然数)との間の類似度 $S(M, N)$ を画像類似度計算手段809が計算する。

【0099】手順911では、手順211と同様に、画像基準処理手段810で、 M が1から20までのいずれかの値をとるときに、(2)式が成り立つかどうか調べ、(2)式が成立するとき、次シーンの代表画像を画像基準モードの代表画像とみなす。手順911の終了後、手順902に戻る。

【0100】手順912では、手順212と同様に、映像の再生と映像の圧縮を中止する。ここまでの手順でファイルサーバー814には、複数の映像要約基準、すなわち全シーンモード、時間判定モード、ハイブリッドモード、画像基準モードによりそれぞれ抜き出された要約情報、すなわち代表画像と先頭フレーム番号が保存された。以下の手順913から手順918までは映像表示装置815において行われる処理である。

【0101】手順913から手順916まででは、映像の要約情報である代表画像を所定の枚数にまで取捨選択する。以下では、映像要約装置で抜き出す代表画像の枚数が24枚に設定されている場合について述べる。

【0102】手順913では、画像判定手段811が、あらかじめ決めてある優先順位に基づいて、どのモードによる代表画像を抜き出すかを決定する。以下では、優先順位の最も高いモードを画像基準モードとし、2番目に高いモードを時間判定モードとし、3番目に高いモー

ドをハイブリッドモードとし、4番目に高いモードを全シーン表示モードとした場合について述べる。なお、この優先順位は、使用者が画像判定手段811に命令を送ることのできるインターフェース手段を設けて自由に設定する態様をとることができる。

【0103】まずモードを優先順位の最も高い画像基準モードにより代表画像を決定する。画像基準モードにより代表画像が決定された後に手順916の分岐処理によりこの手順913に戻って来た場合には、前回決定したモードよりも優先順位が一つ低いモードにより代表画像を決定する。

【0104】手順914では、画像選択の優先順位にもとづいて、手順913で決定したモードの代表画像に順番を付ける。例えば、手順913で決定した代表画像の中で、時間長の長いシーンの代表画像から順番に若い番号をつけていく。同じモードの次の代表画像までの長さが長い代表画像から順番に若い番号を付けてもよい。

【0105】手順915では、若い番号から順番に、映像要約装置で抜き出す代表画像とみなしていく。ただし、映像要約装置で抜き出す代表画像の枚数が24枚を

【0106】手順916は分岐処理である。すでに映像要約装置で抜き出すことに決定している代表画像の枚数が24枚未満のときは、手順913に戻る。そうでなければ、手順917に進む。

【0107】なお、以上の説明では、映像要約装置で抜き出す代表画像の枚数を24枚に設定したが、必ずしも24枚でなくてもよい。また、全てのモードによっても代表画像の枚数が24枚未満のときは手順916で無限ループに入るおそれがある。したがって、全てのモードを選択した後もシーンの数が24枚未満になるときには、その時点での代表画像の枚数をシーンの数と同じになるようにしてもよい。

【0108】手順917では、使用者が見たい部分を効率よく探せるように、システムが映像表示装置上に映像の要約を表示する。例えば、手順915で決定した代表画像を一覧表示する。

【0109】手順918では、使用者が見たい部分を指定する。例えば、マウスなどのポインティングデバイスを用いて、見たい部分の代表画像を指定する。映像表示装置がファイルサーバーから映像データを受け取り、指定された部分から映像を再生する。

【0110】なお、以上の説明では、手順917で代表画像を一覧表示したが、実施の形態1の手順214で述べたように、要約映像再生手段を用いて要約映像を表示してもよい。また、代表画像を一覧表示する画像表示手段を設けて、映像要約装置803の中に組み込んでよい。この場合には、手順917でシステムが画像表示手段を用いて映像の要約を表示する。

【0111】以上の説明では、設定された枚数以下の代

表画像のみを一覧表示しているが、時間判定モード、ハイブリッドモード、画像基準モードの各モードの代表画像の一覧表示を選択できるようにしてもよい。

【0112】以下では、本実施の形態の映像要約システムにおける手順907において時系列のシーンをグループ化することのできる共通画素法によるグループ化について述べる。共通画素法は出願人が先に特願平7-46970号において開示したものである。

【0113】共通画素法は、シーンに共通する色に着目して、砂浜で撮影したシーンが続く場合のような類似背景のシーンまたは様々な人物のバストショット（人物の胸から上が映っているシーン）が続く場合のような、類似被写体のシーンが時系列に連続する場合を検出し、一つのグループに統合する方法である。

【0114】シーンは内容の最小単位である。従って、1つのシーン内のフレーム画像は「同一人物が登場する」などの共通した特徴をもつ。そこで、各シーンの先頭部分の動画画像がシーンを代表するとみなし、この動画画像を代表時空間画像という。

【0115】「色が共通する画素」を同一物体とみなすと、「異なる動きをする同一色の物体」のシーンがグループ化される問題がある。この問題点を解決するために、同一色の物体が、異なる2つのシーンで共に静止している場合と、異なる2つのシーンで共に動いている場合に限って、同一色の物体を同一物体とみなす。即ち、「動きの有無と色が共通する画素」を同一物体とみなす。また、同一グループ内のシーンが共通色比率条件、すなわち、各シーンの代表時空間画像において、「グループ内のシーンの代表時空間画像に共通して現れる同一物体」の画素の総数を全画素数で正規化した値がしきい値 θ_{SHOT} 以上になるという条件を満たすと仮定する。

【0116】図11は共通色比率条件の様子を示す説明図である。図11に示すように、シーン $SC_1 \sim SC_4$ の代表時空間画像に共通して現れる物体がA（背景）の部分のみであり、Aの部分の画素数の占める割合がシーン SC_1 以外でしきい値 θ_{SHOT} 以上になるとき、シーン $SC_1 \sim SC_4$ は共通色比率条件を満足するが、シーン $SC_1 \sim SC_4$ は共通色比率条件を満足しない。 $SC_3 \sim SC_4$ が共通色比率条件を満足するかどうかは不定である。図11の例では、 $SC_3 \sim SC_4$ のA（背景）とB（長方形の物体）の部分が「シーン $SC_3 \sim SC_4$ の代表時空間画像に共通して現れる同一物体」となり、その画素数の占める割合がシーン SC_3 、 SC_4 の両方において100%になるので、シーン $SC_3 \sim SC_4$ は共通色比率条件を満足する。

【0117】このとき、共通色比率条件を満足するシーンを同一グループとみなすと、シーン $SC_1 \sim SC_3$ とシーン $SC_3 \sim SC_4$ が同一グループとなるので、シーン $SC_1 \sim SC_4$ が同一グループとなるはずであるが、シーン $SC_1 \sim SC_4$ は共通色比率条件を満足せず矛盾が生じ

る。従って、共通色比率条件を満足しても、同一グループ内のシーンとは限らない。

【0118】共通画素法では、同一グループ内で隣り合うシーンの間の類似度が、異なるグループのシーンの間の類似度に比べて大きい値になると仮定し、以下の手順1~2でグループの境界を求める。ただし、 $M=1$ とし、 L の初期値は1とする。

【0119】手順1では、シーン $SC_M \sim SC_{M+L}$ が共通色比率条件を満足するかどうか判定する。満足する場合には、 L に1を加えながら、共通色比率条件を満足しなくなるまで判定を繰り返す。図11の例では、 $L=4$ になった時点で手順2に進む。

【0120】手順2では、共通色比率条件を満足する各シーン $SC_M \sim SC_{M+L-1}$ に対して、次シーンとの類似度を求め、類似度が最小になる部分をグループの境界とみなす。図12はグループの境界の決定方法を示す説明図である。同図において、 SC_2 と SC_3 の類似度が他の類似度、すなわち SC_1 と SC_2 の類似度および SC_3 と SC_4 の類似度に比べて小さい値なので、 SC_2 と SC_3 の間をグループの境界とみなす。

【0121】なお、代表時空間画像の代わりに代表画像を用いてもよい。このときは、「色が共通する画素」を同一物体とみなせばよい。

【0122】(第3の実施の形態) 第3の実施の形態は、第2の実施の形態において説明した映像要約システムにおいて、どのモードの代表画像から抜き出すかの選択を、第2の実施の形態の手順913のようにあらかじめ決められた優先順位に基づいてモードを決定するのではなく、映像を解析してモードの優先順位を決定することにより行うものである。

【0123】図9の映像要約システムについて、図13に示すフローチャートを用いてその動作を説明する。

【0124】手順1201~手順1212の動作は、図10に示した手順901~手順912の動作と同じであるのでその説明を省略する。

【0125】手順1213では、画像判定手段811がモードの優先順位を決定する。以下では、画像基準モードの代表画像の枚数が5枚以上になるときは、優先順位の最も高いモードを画像基準モードとし、2番目に高いモードを時間判定モードとし、3番目に高いモードをハイブリッドモードとし、4番目に高いモードを全シーン表示モードとし、画像基準モードの代表画像の枚数が5枚未満のときには、優先順位の最も高いモードを時間判定モードとし、2番目に高いモードをハイブリッドモードとし、3番目に高いモードを全シーン表示モードとした例について述べる。なお、ここでは画像基準モードの代表画像が5枚以上であるか否かを調べてモードの優先順位を変えているが、必ずしも5枚に設定する必要はない。この例のようにモードの自動決定基準を定めたのは、画像基準モードの要約映像が5枚以下の場合には要

約する映像にはヘッドラインが含まれていない可能性が高いという推定に基づいている。なお、このようなモードの自動決定基準は、使用者が画像判定手段811に命令を送ることのできるインターフェース手段を設けて自由に設定を変更する態様をとることができる。

【0126】手順1214では、手順1213で決定した優先順位にもとづいて画像判定手段811がモードを決定する。まずモードを優先順位の最も高い画像基準モードにより代表画像を決定する。画像基準モードにより代表画像が決定された後に手順1217の分岐処理によりこの手順1214に戻って来た場合には、前回決定したモードよりも優先順位が一つ低いモードにより代表画像を決定する。

【0127】以下では、映像要約装置で抜き出す代表画像の枚数が24枚に設定されている場合について述べる。

【0128】手順1215では、手順914と同様に、画像選択の優先順位にもとづいて、手順1214で決定したモードの代表画像に順番を付ける。例えば、手順1214で決定した代表画像の中で、時間長の長いシーンの代表画像から順番に若い番号をつけていく。

【0129】手順1216では、手順915と同様に、若い番号から順番に、映像要約装置で抜き出す代表画像とみなしていく。ただし、映像要約装置で抜き出す代表画像の枚数が24枚を超えたら、作業を中断する。

【0130】手順1217では、手順916と同様の分岐処理を実行する。すでに映像要約装置で抜き出すことに決定している代表画像の枚数が24枚未満のときは、手順1214に戻る。そうでなければ、手順1218に進む。

【0131】手順1218では、手順917と同様に、使用者が見たい部分を効率よく探せるように、システムが映像表示装置上に映像の要約を表示する。例えば、手順1216で決定した代表画像を一覧表示する。

【0132】手順1219では、手順918と同様に、使用者が見たい部分を指定する。例えば、マウスなどのポインティングデバイスを用いて、見たい部分の代表画像を指定する。映像表示装置がファイルサーバーから映像データを受け取り、指定された部分から映像を再生する。

【0133】

【発明の効果】以上で説明した本発明は次のような有利な効果を奏するため、一定の映像要約基準のみで画一的に代表画像を選択する従来の映像要約装置や代表画像を映像の長さ等に関係なく選択する従来の映像表示装置と比べて、映像内容の多様性および使用者の好みの多様性に対応することができる。

(1) 請求項1および請求項2に記載した発明に基づく映像要約装置は、例えばニュース番組においてしばしば起こるように、現場のアナウンサーが事件を説明するシ

ーン等のように前後のシーンと同一背景であっても重要な情報を有している可能性の高いシーンが一定時間以上継続する場合には、そのシーンとその直後のシーンのように重要な情報を有している可能性が高いシーンを背景が類似している前後のシーンから取り出すことができる。グループ化の判断時間(本明細書の実施の形態では8秒)より時間長の長いシーンとしては、ニュース番組のアナウンサーのシーン、インタビューのシーン、登場人物による説明のシーン、フリップの出るシーンなどがある。我々の分析によれば、時間長の長いシーンの直後のシーンは、時間長の長いシーンと同様に重要であることが多い。

【0134】例えば、ニュース番組で現場から事件を報告する場合には、「現場に派遣された登場人物が事件の背景を説明するシーン」に続いて、「事件に関連するシーン」の映像が流れる。この場合、「現場に派遣された登場人物が事件の背景を説明するシーン」の画像よりも、「事件に関連するシーン」の画像の方が、事件の内容を的確に表しており、重要である。しかし、色差ヒストグラム相関を用いた代表画像決定方法では、現場という同一背景のシーンになるため、「事件に関連するシーン」と「現場に派遣された登場人物が事件の背景を説明するシーン」が同一グループになり、「事件に関連するシーン」の画像が代表画像にならない問題がある。また、時間判定モードでも、「事件に関連するシーン」が必ずしも8秒以上にならないため、代表画像になるとは限らない。

【0135】一方、ハイブリッドモードでは、8秒以上のシーンとその直後のシーンが代表画像になるので、「事件に関連するシーン」の画像が代表画像として選ばれ、色差ヒストグラム相関を用いた代表画像決定方法と時間判定モードにない効果が得られる。時間長の長いシーンの中で、アナウンサーのシーン、インタビューのシーン、登場人物の説明のシーンにおいては、直後のシーンと一緒に表示されないと、内容がわからないことが多い。色差ヒストグラム相関を用いた代表画像決定方法を用いたとき、時間長の長いシーンの前後のシーンが表示されるケースが多いが、しきい値が不適当だったり、背景が類似する場合には、直後のシーンが表示されない。しかし、ハイブリッドモードでは、確実に直後のシーンが表示される効果がある。従って、ハイブリッドモードは、色差相関値法や時間判定モードと比較すると、かかる種類の番組の要約においては優位性を有する。

(2) 請求項3および請求項4に記載した発明に基づく映像要約装置は、本編の各記事から2〜3シーンの動画像を抜き出して作成されたダイジェストが最初に流れてから本編が流れる番組で、これらのダイジェスト画像を代表画像となるように本編をグループ化することができる。ダイジェスト部分と本編部分と比較して代表画像を決定するからである。すなわち、画像基準モードはかか

る種類の番組で絶大な効果を有する。従って、画像基準モードでは、映像の内容がダイジェスト部分のシーンで代表され、見たい部分を簡単に探すことができる。

【0136】また、映像表示装置と組み合わせて映像要約システムとして使用することにより、見たいシーンを指定すると、本編の中のそのシーン以降の映像が再生させることができ、見たい記事を簡単に再生できるという従来の映像要約システムにはなかった効果を有する。また、画像基準モードにより代表画像がほとんど選択されない場合には、その映像中にはその映像の内容を要約的に紹介する部分(例えば、ニュース等ヘッドライン等)が存在しないことになる。従って、本モードは映像に要約情報が付属しているか否かを判定することができるという効果を有する。

(3) 請求項5、請求項6、請求項7、請求項8、請求項9、および請求項13に記載した発明に基づく映像要約装置は、複数の映像要約基準を備えているため、単一の映像要約基準により映像を要約していた従来の映像要約装置と比較して、映像の内容の多様性により適切に対応することができる。使用者は、番組等によって映像要約基準を変更したり組み合わせたりすることが自由でできる。

(4) 請求項7および請求項8および請求項9に記載した発明に基づく映像要約装置は、使用者が代表画像の枚数の上限を指定できるので、代表画像の枚数を少なく抑えることができ、従来法に比べて見たい部分を簡単に探すことができる。さらに、あらかじめモードの優先順位を決めておいた場合には、使用者が映像の特徴に合わせて代表画像を決定できる。また、画像基準モードの代表画像の枚数がしきい値以上になるかどうかによって代表画像の決定方法を切り替える場合には、ダイジェストが存在する映像で、ダイジェスト部分の画像を優先して抜き出すことができるという効果をも有する。したがって、使用者が要約映像検索する効率が従来の映像要約システムと比べて向上する。

(5) 請求項10に記載した発明に基づく映像表示装置は、映像の情報およびその映像の要約情報を記録しているファイルサーバーの情報を利用して、各代表画像の位置付近の一定時間の映像を次々に見ながら見たい部分を指定できる、すなわち動画と音声を用いて見たい部分を探すことができるので、音声や被写体の動きを考慮しながら見たい部分を指定できるようになる。さらに、重要性の高いシーンなどから代表画像が選ばれているので、再生される映像を見て、使用者が冗長に感じることはない。

(6) 請求項11および請求項12に記載した発明に基づく映像要約装置は、映像の代表画像およびその前後よりなる要約映像を再生する際に、重複して再生する部分が無いようにしながら要約映像を再生するので、冗長な再生にならない。さらに、各代表画像の位置付近の一

10

20

30

40

50

25

定時間の映像を次々に見ながら、見たい部分でこの再生を中断すると、中断した位置のフレーム画像の内容を代表する代表画像を含む複数の代表画像を一覧表示するので、中断した位置付近に存在する「見たい部分に関連する内容」を簡単に探すことができるという効果を有する。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態における映像要約システムのブロック図

【図2】本発明の第1の実施の形態における映像要約システムの動作を示すフローチャート

【図3】シーンチェンジ検出手段がシーンチェンジを検出する方法と前シーンを示す説明図

【図4】画像基準モードにおいて、画面をブロックへ分割することを示す説明図

【図5】本発明の第1の実施の形態における要約映像の作成方法を示す説明図

【図6】本発明の第1の実施の形態における見たい部分の先頭からの再生開始の様子を示す説明図

【図7】本発明の第1の実施の形態におけるユーザーインターフェイス手段を介して見たい部分を指定するための画面を示す説明図

【図8】本発明の第1の実施の形態における要約映像の再生部分を示す説明図

【図9】本発明の第2の実施の形態における映像要約システムのブロック図

【図10】本発明の第2の実施の形態における映像表示システムの動作を示すフローチャート

【図11】本発明の第2の実施の形態における共通色比率条件の様子を示す説明図

【図12】本発明の第2の実施の形態におけるグループ

の境界の決定方法を示す説明図

【図13】本発明の第3の実施の形態における映像表示システムの動作を示すフローチャート

【図14】従来の映像表示装置におけるシーン一覧表示を示す説明図

【図15】従来の映像表示装置における映像要約結果の一覧表示を示す説明図

【図16】従来の映像要約システムのブロック図

【図17】従来の映像要約システムの動作を示すフローチャート

【図18】従来の共通色比率法で検出されるシーンチェンジを示す図

【符号の説明】

101、801 ビデオディスク装置

102、802 VTR

103、803 映像要約装置

104、804 画像取り込み手段

105、805 シーンチェンジ検出手段

106、806 時間判定処理手段

107、807 グループ生成手段

108、808 グループ追加手段

109、809 画像類似度計算手段

110、810 像基準処理手段

811 画像判定手段

111、812 制御装置

112、813 映像圧縮装置

113、814 ファイルサーバー

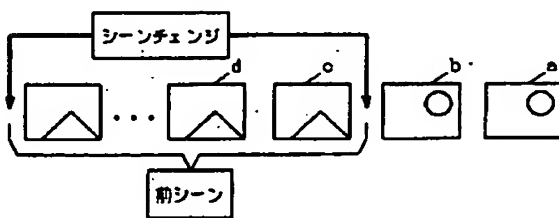
114、815 映像表示装置

115 要約映像再生手段

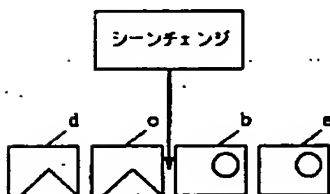
116 ユーザーインターフェイス手段

117 映像再生手段

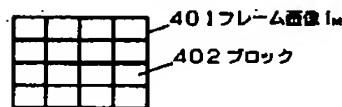
【図3】



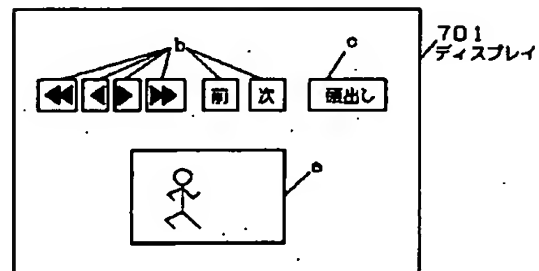
【図18】



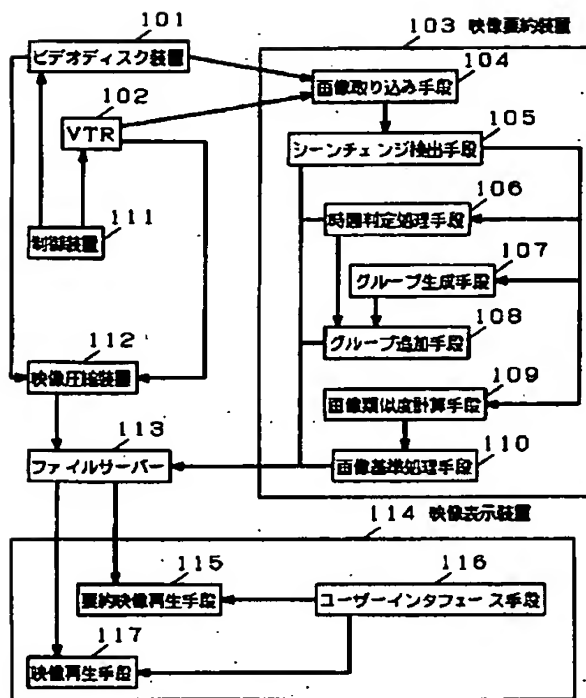
【図4】



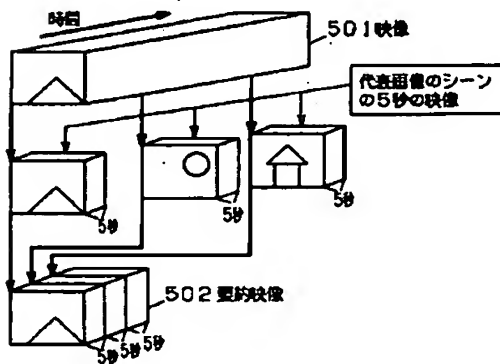
【図7】



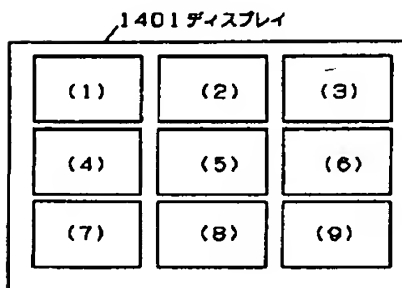
【 図1 】



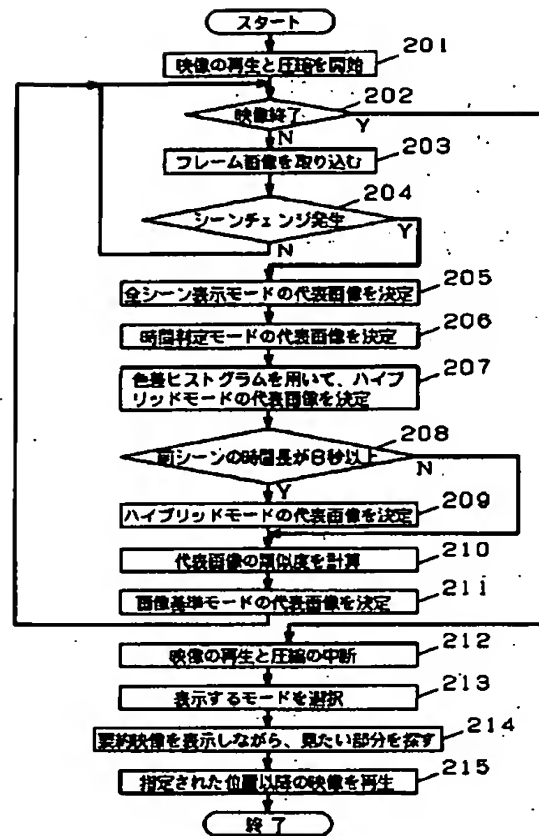
【 図5 】



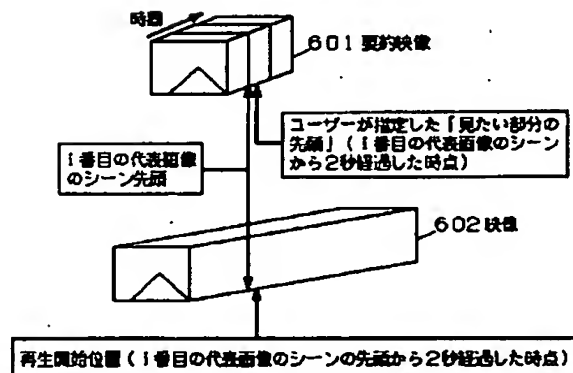
【 図14 】



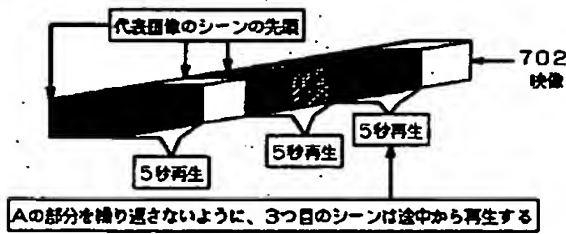
【 図2 】



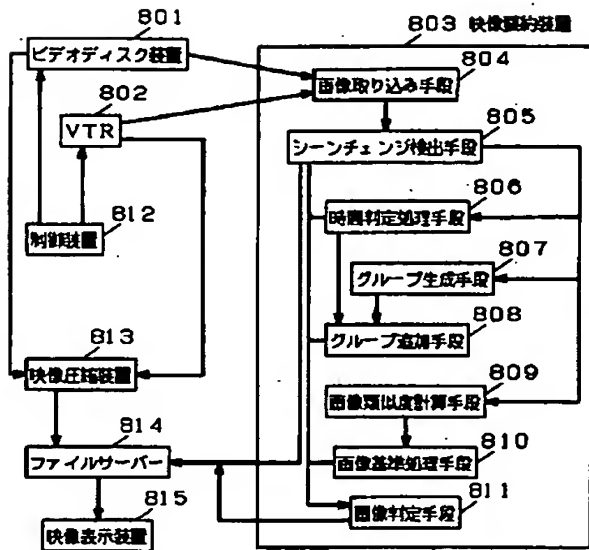
【 図6 】



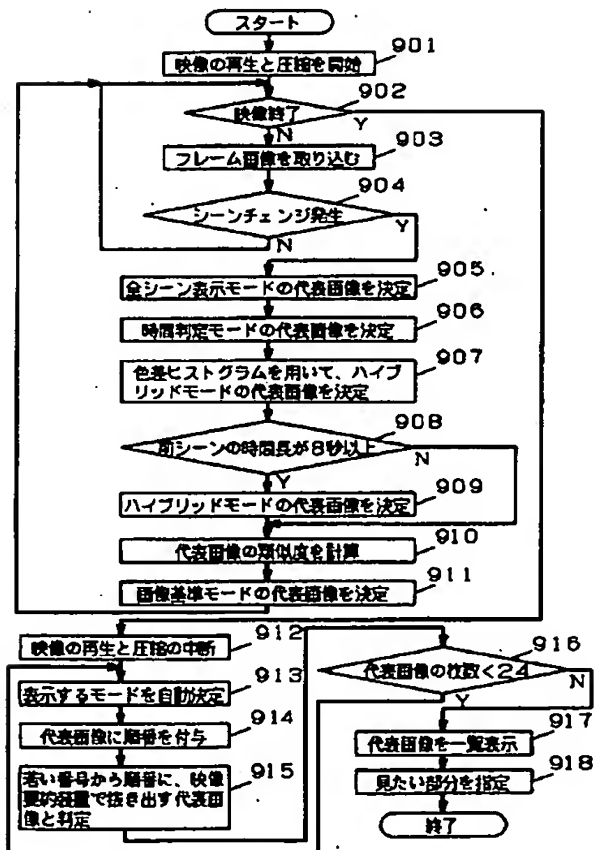
【 図8 】



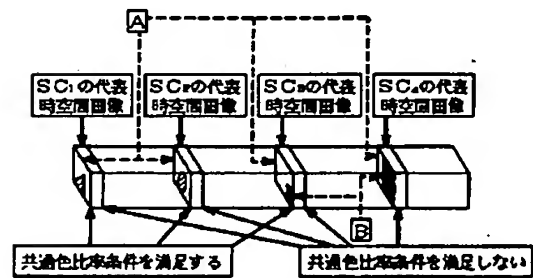
【 図9 】



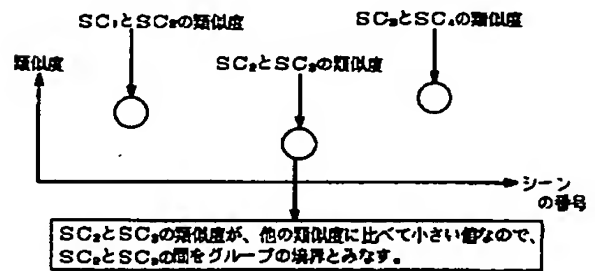
【 図10 】



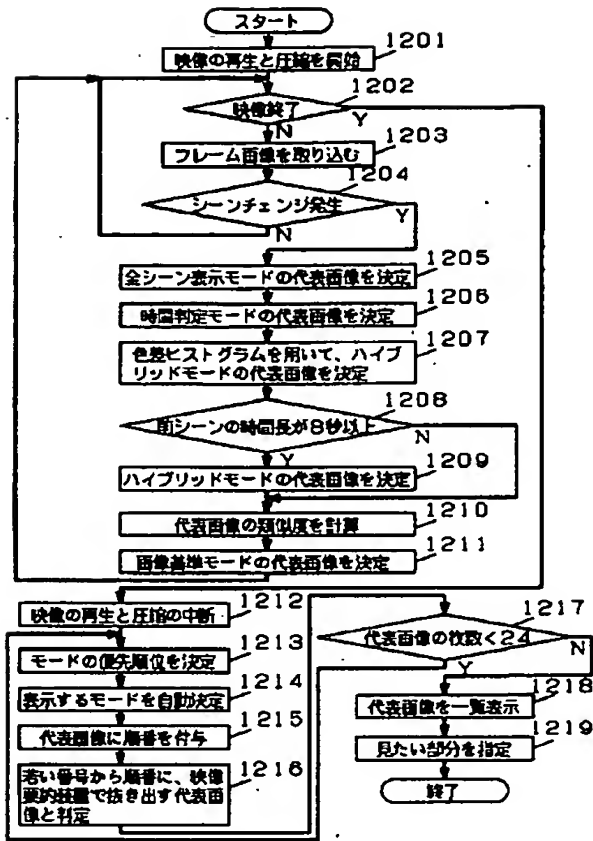
【 図11 】



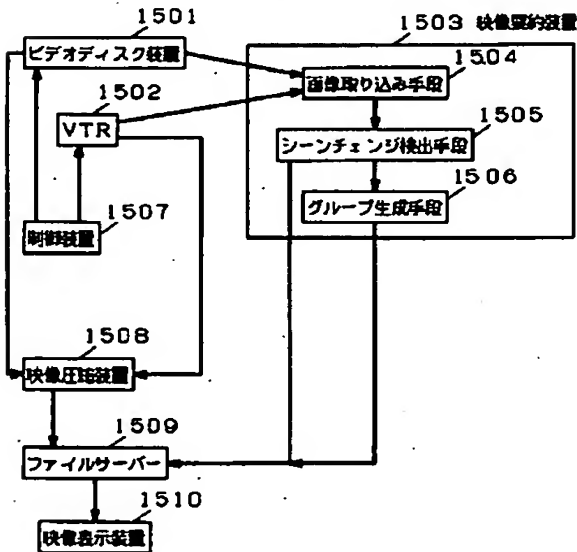
【 図12 】



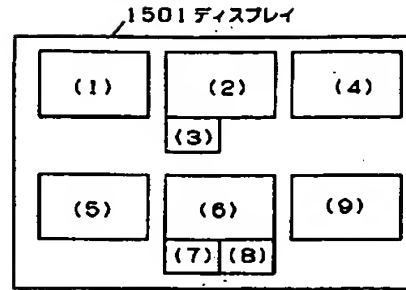
【 図13 】



【 図16 】



【 図15 】



【 図17 】

